Disability, Rationality, and Justice: Disambiguating Adaptive Preferences[1]

Jessica Begon, University of Oxford
jessica.begon@nuffield.ox.ac.uk

This is a slightly longer draft of a chapter forthcoming in: *The Oxford Handbook of Philosophy and Disability*, David Wasserman and Adam Cureton (eds.), (Oxford University Press). If citing, please use the final, published version.

A growing body of evidence suggests that the self-reported welfare levels of disabled individuals are no worse, and sometimes better, than that of non-disabled people. This is so surprising to individuals without disabilities, many of whom consider the prospect of becoming disabled as one of the greatest misfortunes that can befall us, that it has been called the 'disability paradox'.[2] Yet not only do disabled individuals rate their subjective quality of life highly, many also resist the suggestion that their conditions should be cured, and refuse cures when offered; many argue for the right to have disabled children, and against policies of pre-natal screening for various disabilities; they march in Disability Pride parades, and insist that disabled lives are not inherently worse than non-disabled lives. Nonetheless, the idea of disability as tragic remains the common-sense view amongst most able-bodied (and, indeed, some disabled) people.

Faced with such a divergence of views, whom should we trust? On the one hand, it may seem that disabled people are far better placed to understand the realities of living with their conditions and that, as such, it is their testimony we should rely on. Yet, on the other, we may think that disabled individuals' experience is exactly what gives us reason to doubt their testimony. It is widely acknowledged that in conditions of great hardship or deprivation individuals may cope with their circumstances by claiming to prefer and, indeed, coming to prefer, their situation to any alternative. However, such 'adaptive preferences' need not constitute decisive evidence that these conditions are not, in fact, deplorable. We do not, then, consider women's support of patriarchal norms to justify them, or

[1] For helpful comments and discussion on earlier drafts of this paper I would like to thank audiences at the Centre for the Study of Social Justice (CSSJ) Seminar at the University of Oxford, and the CELPA Seminar at the University of Warwick, as well as Cécile Fabre, Carl Fox, Alex Geddes, and Katy Wells, and the editors of this volume, Adam Cureton and David Wasserman.
[2] Albrecht and Devlieger 1999

an abused partner's preference to remain in the relationship to mitigate this abuse. Along similar lines, we may claim that disabled individuals' apparent satisfaction with their lives is not reliable evidence against the claim that disability must be disadvantageous.

This might seem like a neat solution: if disabled individuals' preferences are simply hopelessly deformed by the terrible circumstances in which they were formed, then we can treat their positive claims about disability as suspect, and maintain our intuition that it is a tragedy after all.[3] It is certainly a solution that allows many people to maintain their strongly held intuitions. However, we should be cautious about simply ignoring the preferences and beliefs of a group who are already subject to considerable injustice. The unjustified silencing or mistrust of already underrepresented groups may constitute a serious epistemic injustice, quite apart from the possibility of being used to justify interference in their lives and choices. There has, on this basis, been a considerable backlash against the tendency amongst Western feminists' to see women in the third-world as the 'dupes of patriarchy', whose preferences can be ignored.[4] Elizabeth Barnes, too, has argued against adaptive preference models of disabled individuals' preferences on the basis that when "misapplied…[they] can simply entrench pre-existing biases"[5].

However, such worries about its misapplication should not lead us to abandon the concept of adaptive preferences entirely. If we were to simply take individuals' claims to be satisfied with their lot at face value, this may mean that mistreatment, injustice, and oppression would not be recognised or rectified. If those subject to oppression do not identify their own mistreatment, they may not be entitled to redress. Indeed, it may not be considered mistreatment at all, since 'no one complains'. Thus, some balance needs to be struck between taking individuals' positive assessment of their quality of life and their preferences for their current way of being as definitive, and ignoring them entirely.

A further problem with diagnosing adaptive preferences, less frequently acknowledged, is the vagueness of the term, and the breadth of the phenomena it

[3] My focus is cases where merely preferring a life with an impairment is thought to cast doubt on the reliability of an individual's testimony. Thus, I am concerned with physical and cognitive impairments that are not considered, in themselves, to undermine an individual's capacity for autonomous agency, and so do not consider the issues raised by severe cognitive impairment (nor take a view on which should be considered 'severe').
[4] Narayan 2002: 418. Also, Khader 2011, 2012, 2013; Jaggar 2005.
[5] Barnes 2016: 137

can be used to cover. From the opening paragraph we can see the various experiences and attitudes that can be roughly grouped – and rejected – as adaptive preferences, despite the fact they cannot all be plausibly interpreted as preferences. Further, we might reasonably think that if our concern is adaptation, then *all* preferences formed in response to our circumstances should also be included. It should be clarified from the outset, then, that 'adaptive preferences' as the term is used in the literature (hereafter APs) is both broader and narrower than an intuitive understanding of the concept. Broader because it includes not just preferences, but something like 'evaluative states'[6]; and narrower because it includes only those adaptations that are, in some sense, unreliable. However, even accepting these restrictions, ambiguity remains: in particular, concerning what constitutes *relevant* unreliability.

My goal in this paper is to respond to the problems of both misapplication and vagueness. By clarifying the various phenomena loosely categorised as APs I show that correctly diagnosing an individual's preference as adaptive need not constitute epistemic injustice, nor insultingly malign their capacities as a rational agent, and further, that our account of relevantly unreliable APs should be sensitive to our particular goals and context. Thus, no single account of APs will be satisfactory for all the purposes to which it is put. Our diagnosis of APs will depend both on the role these preferences play (for example, being used to fill in the contents of a theory of justice, or to generate an account of well-being), and on how information about preferences is extracted (from individuals' choices, expressed satisfaction, or expressed counterfactual preferences). Thus, an individual may, for example, be a reliable guide to their well-being, but not a reliable guide to what they are owed as a matter of distributive justice (and *vice versa*); they may be a reliable guide to the content of their distributive entitlements, but not to whether they have had these entitlements met; and their choices and apparent satisfaction may be unreliable in all these regards, whilst their expressed counterfactual preferences are not. It would be a mistake, then, to distinguish individuals whose preferences are adaptive from those who are not; and a mistake, too, to distinguish preferences that are adaptive (for all purposes, in all contexts) from those that are not.

In particular, I defend a distinction between 'well-being adaptive preferences' (WBAPs), and 'justice adaptive preferences' (JAPs). WBAPs are APs as classically

[6] As Terlazzo (2015: 1) notes, the term incorporates both "comparative preferences proper" and "non-comparative states like desires, judgements of values, etc.".

understood: individuals adapt to constrained options in ways that render their resulting preferences procedurally irrational and non-autonomous, and so undermine their possessor's authority about their own well-being. In the context of social justice, however, this account is both over and under-inclusive: it includes preferences that do not warrant state interference, and excludes preferences that are a rational response to our circumstances, yet should not be allowed to inform a theory of justice. Thus, I contend that we also need to identify JAPs: preferences that are a poor guide to individuals' entitlements. Some of these are rational (and so not also WBAPs), whilst others are both justice and well-being adaptive. JAPs and WBAPs, then, are overlapping but distinct categories. I elaborate on my understanding of WBAPs (§1) and JAPs (§2), and argue that both concepts are needed to understand the different ways in which individuals' preferences may be unreliably adaptive, and the different responses that may be appropriate.

Whilst, I argue, not all APs are irrational they *are* all non-autonomous. Thus, we may worry that the account remains vulnerable to the objection that diagnosing APs has the insulting implication that individuals are defective agents or 'dupes'. However, this is not the case. First, in many cases APs are a rational response to unjust circumstances. In these instances, autonomy is undermined by the limitation of circumstances rather than defects in individuals' capacities. Second, all (or almost all) of us possess some WBAPs: preferences we hold for reasons that are necessarily opaque to us. That some of these are relevant to justice (JAPs) – and thus, are justifiably ignored in formulating a theory of justice, and may legitimate state interference – is again, I argue, not the result of faulty individual capacities, but of circumstances. Thus in clarifying the concept of APs, I will also demonstrate that being diagnosed as having an AP need not have the insulting implications ordinarily supposed.

1. Autonomy, Proceduralism, and Well-Being Adaptation

1.1 Adaptation and the Political Project

Classic accounts of APs – developed by Jon Elster and his critics – are modelled on Aesop and La Fontaine's 'Fox and Grapes' parable.[7] On realising he cannot reach the grapes he desires, the fox insists 'grapes are too sour for foxes', and he

[7] For example, Elster 1987; Bovens 2002; Colburn 2011; Barnes 2009.

did not want them anyway. It is assumed this response is irrational, and not a reflection of his best interests. These accounts root this irrationality in *procedural* flaws in the process of preference formation. They focus on capturing the idea – central to our intuitions about the fable – that the fox's preference is unreliable because he is unconsciously 'fooling himself': he has failed to acknowledge his limitations, or recognise that the *real* reason he no longer prefers grapes does not concern their sourness. If, instead, the fox's response to being unable to reach the grapes were to consciously cultivate a preference for a sweeter, lower-hanging fruit, this would not be an AP, but conscious, rational, and autonomous *character planning*.

When the problem of APs was raised in the context of social justice (notably development ethics) it might have been natural to assume that the APs referred to were of this procedurally-flawed, 'sour grapes' form – especially given the lack of a clear alternative account. Yet examples used in the literature belie this: "the hopeless beggar…the dominated housewife, the hardened unemployed…[who] take pleasures in the small mercies and manage to suppress intense suffering for the necessity of continued survival"[8] are not obviously 'fooling themselves'. Indeed, in many cases individuals seem well aware of their suffering, and make considerable conscious effort to suppress it. Moreover, there are cases where preferences *are* unconsidered, yet do not seem to be a matter of social justice, nor to warrant public mistrust or political interference. Thus, we may mistrust a woman's unconsidered preference to undergo female genital cutting (FGC), but not a similarly unconsidered preference *not* to undergo FGC. Indeed, we may mistrust a choice to undergo FGC that *is* procedurally autonomous and rational, if it is made against unjust background conditions – for example, as a means to marriage, which in turn is necessary for economic security and social status.

We may, therefore, be tempted to conclude that when thinking about APs in the context of social justice, these procedural accounts are useless. Indeed, this is the conclusion those writing in this context have come to. Serene Khader, for example, has argued against such accounts on the basis that they will lead us to mistrust preferences that are not "worthy of public interrogation"[9], thus generating conclusions inconsistent with our intuitions.[10] Drawing on Khader's

[8] Sen 1988: 45-46. See Khader (2011: 8-10; 42-46) for discussion of the discrepancy between Sen's examples and Elster's account.
[9] Khader 2011: 75
[10] Khader 2011: 74-106; Khader 2009

work, Rosa Terlazzo argues that all accounts of APs must meet the following criterion of 'political efficacy':

> An account must provide us with a valuable tool for combatting marginalization and oppression, by explaining which preferences deserve social suspicion and why they do so. It must include paradigm cases of adaptive preference and exclude preferences that have nothing to do with the political project.[11]

It is true that traditional accounts will fail this test. This is unsurprising given that they were never designed to pass it, concerned as they are with procedural rationality rather than social justice. It is also true that we need an account of APs that is politically efficacious – for which reason I offer an account of JAPs. However, it does not follow that there is no value in maintaining an account of APs as procedurally non-autonomous and irrational, and so a poor guide to individuals' interests – that is, of WBAPs.

1.2 Well-Being Adaptive Preferences

How, then, might an account of procedurally non-autonomous preferences be cashed out? Elster saw the distinctiveness of APs in the fact that they are formed unconsciously, or 'behind our backs', in light of diminished options.[12] Nussbaum (amongst others) has argued convincingly that Elster's account is likely to be over-inclusive, demanding conditions of procedural rationality few of our preferences meet. As she points out, we are not – and should not – be "suspicious of any desire that is formed through [unconscious] adjustment to reality"[13]. This is not simply the criticism that this account cannot pick out APs relevant to social justice, but the broader claim that it cannot distinguish unreliable preferences in any context. Too many of our preferences are unconsciously formed for this account to be useful: "[t]he desire for food is not normally induced through conscious hunger creation"[14]. This is true, too, of more significant preferences: for example, abandoning our childhood career aspirations, even if we are unconscious of the reasons, is not usually thought to cast doubt on the rationality or reliability of our commitment to our adult occupation.

However, it may be that Elster is really concerned with APs not because they are unconscious, but because they lack autonomy. If this is so, then the absence of a

[11] Terlazzo 2015: 5. Also, see Nussbaum 2000; 2001; Barnes 2016; 2009.
[12] Elster 1987: 117-119
[13] Nussbaum 2001: 78
[14] Colburn 2011: 57

clear account of autonomy is a serious omission (and one later authors have tried to fill).[15] I focus here on Ben Colburn's account, which seems to best capture the cases we have in mind. In brief, Colburn argues that APs are the result of 'covert influences'; that is, the reasons for our APs are necessarily hidden from us, such that they could not explain our commitment if we were conscious of them – we would either adduce independent reasons for our preference or repudiate it. Thus, the fox's claim that grapes are sour could no longer explain his preference not to eat them if he were aware that the real (and covert) cause of this preference was their inaccessibility.

This does not imply all unconsciously-formed preferences are adaptive (as Elster's account seems to), nor that any influence on our preferences is suspect. Rather, we should be suspicious of APs because they result from influences their possessor does not, and would not, reflectively endorse, given their current disposition and convictions. To possess such APs, then, is to lack procedural rationality, insofar as we are acting for reasons we could not accept. This also undermines an agent's independence, since someone cannot be "deciding *for herself*"[16] when acting on reasons necessarily hidden from her. If, for example, we discover that the reason we have revised our career aspirations is that we lack the talent to pursue our original dream, and possess the aptitude for our current path, this need not cause us to doubt our goals. Yet imagine if we realised we only abandoned our dream of being a scientist due to being subtly undermined by a sexist science teacher. We do not endorse his view (that 'women are not cut out for science'), so this could not explain our preference. In this case we *may* question the autonomy of our apparent preferences, and our well-being may improve if we revised them.[17]

We may worry that such an account is under-inclusive when applied to the political task, since many paradigm APs seem to be consistent in this way. For example, a woman may internalise a complete set of sexist and patriarchal norms, such that her acceptance of a lack of economic opportunities and rights to political participation are entirely consistent with her more general views about women's capacities and proper role within the home and society. She would,

[15] Colburn 2011; 2010; Bovens 2002; Zimmerman 2003.
[16] Colburn 2010: 26
[17] Complications arise when we have pursued a plan of life for a sustained time, since our well-being may not be improved by abandoning it, even if the reasons we originally engaged in it were not ones we would endorse. If we can give independent reasons for our current preference it ceases to be a WBAP (we now see more value in philosophy than science); if we cannot, it remains a WBAP that we would be better-off for repudiating. I return to this case below (§2.3).

then, endorse the reasons for her preference: they could function as the explanation for her commitment even if no longer hidden from her. Disabled individuals' preferences, too, are often consistent in this way, so this approach cannot be used to discount them[18] (though given that the reliability of disabled individuals' preferences is what is at issue, it is not clear whether this should count against an account of APs).

In response, we can point out, first, that our goal here is not political: we are attempting to identify procedurally non-autonomous and irrational preferences, which are an unreliable guide to the needs of those who have them. It is JAPs that warrant public suspicion, scrutiny, or even intervention. Thus, accepting that individuals who have internalised a complete set of oppressive norms may be considered rational, and a reliable guide to their interests, does not imply they are necessarily immune to such scrutiny or intervention, as §2 will discuss.

In some cases we can distinguish individuals' 'adaptive *choices*' from their counterfactual preferences. For example, a disabled individual may choose not to pursue higher education on the basis that the institutions in which it is offered tend to be exclusive and unaccommodating. This may be consistent with their further preferences (not wishing to associate with people or institutions that mistreat them), and being aware of their reasons need not lead them to change their mind. However, they may prefer a situation in which accessible higher education was available. In such a case, I contend, neither the choice nor the preference should be diagnosed as a WBAP (both are rational) – yet, as §2.3 will argue, the choice (though not the preference) is *justice* adaptive.

There are also cases in which individuals' reasons to acquiesce in their mistreatment are more deep-rooted, and are not accompanied by counterfactual preferences for a more just alternative. For example, if the above individual has internalised a conception of disabled people as 'not worthy' of higher education, they may not prefer it to be available in an accessible form since they simply think 'it's not for people like them'. Further, the internalisation may be sufficiently complete that this view coheres with their other beliefs and preferences, and would not change under conscious examination. It may be objected that this is psychologically implausible – experience of oneself and others as competent and intelligent will necessarily be incongruent with such a

[18] Barnes 2016: 129

conception of disabled individuals. Nonetheless, it does seem possible that sustained oppression could lead to a fully consistent conception of oneself as worthless, at least with regards to certain opportunities.[19]

Yet, in this case too, we should not diagnose this as a WBAP. The individual is not necessarily unconscious of the reason for their commitment, and we have no reason to doubt their authority regarding their well-being *given their current preference set*.[20] However, as in the previous case, the fact that this is not a WBAP does not mean this is not a JAP: though this preference may be *rational*, and a good guide to their current interests, this does not mean it is fully *autonomous*, or a reliable guide to their distributive entitlements. Thus, allowing that individuals may be rational even when their self-conception is seriously undermined does not mean that their acquiescence will be deemed autonomous, nor their underlying oppression legitimised.

Before finishing this section it is worth briefly considering how my account of WBAPs relates to Barnes's 'discordance view', since hers is one of the few accounts of APs developed directly in relation to disability. Barnes argues that preferences are unreliably adaptive if there is "a clear disconnect" with "the rest of the person's life", including "what produces things like anxiety, fear, and antipathy in their lives".[21] Barnes does not flesh out this approach, so it is difficult to get a clear sense of how it would apply in practice, yet it seems unlikely to capture all the cases we would want. A battered woman might, as Barnes points out, experience anxiety, fear and depression. However, women who have internalised more mundane forms of sexist oppression may not experience such cognitive dissonance, yet this surely would not give us reason to overlook it. Further, situations and conditions we on-the-whole prefer can cause stress, anxiety and fear, and such preferences are not obviously unreliable. Barnes primary concern is the epistemic injustice entailed by unwarranted attributions of APs. Avoiding misdiagnosis is certainly important but, as it stands, her positive proposal provides little guidance as to when AP diagnoses *are* warranted.

[19] See Khader (2011: 13-17; 2012: 307) on partial and complete losses of self-worth.
[20] For discussion of similar cases, see Baber 2007: 199-120.
[21] Barnes 2016: 140

1.3 Why Identify WBAPs?

To summarise, WBAPs lack procedural autonomy and rationality since they are the result of influences that agents are necessarily unconscious of.[22] Consequently, they are an unreliable guide to individuals' interests. In the context of social justice this concept is over- and under-exclusive. First, it *excludes* preferences that are rational, and so not WBAPs, but which may be a poor guide to our distributive entitlements, and should be overlooked for the purposes of justice.[23] For example, the preferences of disabled individuals not to enter exclusionary institutions of higher education may be rational, and a good guide to their interests in these unjust circumstances, but their satisfaction should not imply justice has been done.

Second, an account of WBAPs *includes* preferences that are irrelevant to justice. The fox, for example, has a WBAP, but this does not seem to be an injustice the state should rectify. As a more concrete case, consider preferences for particular aesthetic experiences. Deaf individuals, for example, may come to prefer a life without music as a result of influences that they would not reflectively endorse.[24] I assume that whilst justice requires that individuals have certain opportunities – maybe even opportunities for aesthetic experiences – it does not require that individuals have every possible opportunity in such a domain, or even the opportunities they would most prefer (assuming scarcity of resources). Thus, if deaf individuals have access to other forms of aesthetic experience, their inability to hear music is not a concern of justice, so nor is it a concern of justice if their preference not to listen to music is a WBAP.

Given that an account of WBAPs fails to pick out cases of adaptation that seem relevant to justice, we may wonder whether we should follow in the footsteps of

[22] What does it means to be unconscious of the reasons for our preferences? Imagine, for example, a disabled individual who learns, perhaps in therapy, that they have some preference as a result of being subject to stigma. Yet despite 'knowing' the primary cause of their preference they do not attend to it, push it to the back of their mind, and deceive themselves into thinking they have good reasons for this preference. There will always be borderline cases. However, insofar as someone continues to deceive themselves about the reason for their preference, and fails to take on board or internalise the additional evidence they now have, they are not truly conscious of this influence. They are in some sense aware of the pernicious influence (the stigma); yet continue to insist that the real reason for the preference is something else ('I didn't want to go to university anyway'). Thus, this remains a WBAP. (Thanks to Adam Cureton and David Wasserman for pushing me on this point.)

[23] I suggest that there are three categories of such preferences: non-autonomous character planning, justice adaptive choices, and the coherent internalisation of oppressive norms (§2.3).

[24] This is likely true of many individuals' tastes, and there is no reason to think that those with impairments are particularly vulnerable to such covert influences.

other theorists working on APs in the context of social justice, and abandon any attempt to identify such preferences. However, first, the arguments for doing so tend to be based on the assumption that diagnosing a WBAP will wrongly imply that its possessor is unusually irrational or unreflective, or that their preferences warrant public concern or interference. Khader, for example, objects to autonomy-based accounts on the basis that they entail the "dangerous and unwarranted empirical assumption…that people with APs reflect on their behaviour less than everyone else"[25]. However, since WBAPs are not intended to meet the criteria of political efficacy, they need not capture only paradigm cases, where interference seems justified. On the contrary, we can acknowledge that many people's preferences, much of the time, are WBAPs.[26] Adopting an account of WBAPs, then, does not require reaching the dubious conclusion that only oppressed and mistreated individuals have them, nor that they are less reflective or worse at reasoning, nor that possessing WBAPs justifies interference.

Furthermore, there are positive reasons to adopt such an account. Most obviously, to achieve conceptual clarity. If, as I contend, the term 'APs' is used to capture two distinct phenomena – wherein those discussing APs in the context of social justice have in mind something quite different from those discussing APs in the context of autonomy and rational choice – it is worth making this explicit. There are also more practical benefits to distinguishing WBAPs and JAPs. JAPs need not also be WBAPs but the two categories do overlap and, as §2.4 considers, we should treat *rational* JAPs differently from JAPs that are also well-being adaptive.[27] Drawing this distinction also makes clear that having a JAP need not imply we are irrational, and that *not* having any JAPs does not necessarily mean we are fully rational and autonomous.

Further, identifying WBAPs that are not JAPs is useful given that domains of justice might change, and WBAPs that were once irrelevant from the point of view of justice might cease to be so. In the past, ensuring access to aesthetic experiences and leisure may have been considered to be beyond the scope of government activity – in future, other things might be included. Finally, whilst WBAPs may not warrant direct state action, acknowledging the irrationality of

[25] Khader 2011: 80

[26] However, whilst everyone has WBAPs, individuals who are subject to oppression are more likely to have JAPs *as well* (as §2 will discuss).

[27] Specifically, the former case may simply require an alteration of circumstances so that an individual's choices can reflect their preferences, whilst the latter may require deliberation, discussion, and consciousness-raising, alongside the provision of options.

many of our preferences may guide individual behaviour: encouraging us to critically reflect on our preferences, and, if we cannot find reasons for them we would reflectively endorse, change them. Further, though it may be unjustifiably paternalist for a state to interfere in an individual's WBAP, if they are widely prevalent the state *may* legitimately attempt to reduce this prevalence – for example, with measures that promote individuals' capacity for reflection.[28]

2. Perfectionism, Substantive Accounts, and Justice Adaptation

2.1 The Need for a Substantive Approach

If our goal is to combat injustice, and ensure that preferences that endorse and perpetuate oppression are disregarded, then a more substantive account of APs is needed. At the far end of this spectrum is Nussbaum's account, according to which APs are simply those with the wrong content: preferences formed in light of diminished options for what we ought not to prefer.[29] Hence, for Nussbaum, identifying APs requires "a substantive theory of justice and central goods"[30]. Whether a preference is rational or 'considered' is not decisive in determining whether it is reliable. Thus, if an unconsidered preference is for a good way of life (say, economic empowerment) and a considered preference is not (physical abuse), it is the former that is reliable.[31]

Nussbaum's perfectionist approach will certainly allow us to identify paradigm cases, and provide a tool for combatting oppression. We may worry, however, that this approach will be overly narrow, including as reliable only preferences for what has been predetermined to be valuable. Thus, "persons whose conceptions of the good reject the items on the perfectionist list will not be shown the respect granted to those with 'proper' preferences"[32]. Particularly worrying in the current context, it may lead to the exclusion of disabled individuals' preferences if disability is taken to be sub-optimal. Indeed, Barnes has argued that since "[p]hysical disability represents, according to the

[28] This seems to be required by Colburn's 'autonomy-minded liberalism' (Colburn 2010: 94-98).
[29] Nussbaum 2001; 2000: 122-142. On some interpretations, this is Sen's view too (e.g. Barnes 2016; 2009a). However, Sen's approach is complex and not always clearly defined, and I will not take a stand on how it should best be understood here (see Qizilbash 2007).
[30] Nussbaum 2001: 79
[31] E.g. Nussbaum 2001: 84. Given her appeal to a substantive account, it may seem that Nussbaum is not 'relying' on preferences at all. Indeed, they primarily play a heuristic role in her account (Nussbaum 2004: 200; Begon 2015).
[32] Terlazzo 2014: 186

capabilities approach, an absence of one or more basic goods (bodily integrity, physical health etc…)"[33], then it "*cannot* be as optimal as a relevantly similar non-disabled life"[34]. Thus, disabled individuals' preferences for their life cannot be considered reliable. As I have argued elsewhere, it is not clear that the capability approach needs to evaluate disability so negatively.[35] Nonetheless, the general worry remains that Nussbaum's approach excludes those who dispute her conception of value, and "counts a number of widely and sincerely-held conceptions of the good as necessarily non-autonomous"[36].

In light of these problems, both Khader and Terlazzo offer substantive proceduralist accounts of APs, both devised for the development context, and focussing on the preferences of women in the developing world. Both set substantive conditions on preference formation, but aim to avoid ruling out any preferences on the basis of their content alone. Thus, for Khader, "APs are preferences incompatible with an agent's basic wellbeing…formed under unjust conditions"[37]. Khader, then, is more willing to focus on content, seeing APs as any "behavior or belief whereby an individual perpetuates her deprivation", though this must be "causally related to her deprivation", in the sense that it "would disappear upon exposure to superior conditions and/or information"[38]. As such, we have special reason to worry about preferences with specific content – those that perpetuate deprivation, independently and substantively defined – but they are only APs if they arose in particular circumstances. Preferences formed in the same circumstances but with different content are not APs (an oppressed woman's preference for pineapple over mango); nor are preferences "nonconducive to basic flourishing" [39] formed in conditions conducive to flourishing (the decision to endanger one's life by engaging in extreme sports).

Terlazzo argues that individuals should engage in critical reflection on their preferences, and do so in the presence of recognised alternatives that should be both valuable and 'live', in the sense that individuals can reasonably see

[33] Barnes 2009: 2
[34] Barnes 2009: 6
[35] Begon 2015. It is possible, though, that this is an implication of Nussbaum's account (e.g. Nussbaum 2006: 155-223).
[36] Terlazzo 2015: 9
[37] Khader 2012: 302
[38] Khader 2013: 313
[39] Khader 2011: 17

themselves exercising them given their current values and ambitions.[40] The substantive conditions, then, concern the specification of these valuable live options. Further, to meet the criteria of political efficacy she argues that we should restrict the category of APs to "*core* preferences: that is, preferences that are centrally important to a person's broader plan of life or sense of self"[41]. Terlazzo is particularly concerned to remain neutral about the content of APs, but whilst she does not necessarily rule out any conceptions of the good, she does contend that a "unifying characteristic" of all APs will be that they are ones that people would be "highly unlikely to want to have were they to reflect on their own preferences in a considered way and in better circumstances"[42]. This may suggest an implicit judgement about their *likely* content.

Both Khader and Terlazzo emphasise that having APs need not be irrational, or imply an agent lacks the capacity for autonomy. Khader delineates the different forms of AP.[43] She allows that individuals occasionally have 'paradigmatic APs', wherein someone "perpetuates injustice against herself because of a near-completely distorted worldview"[44]. More often, though, individuals are mistaken in some domains, but do not entirely lack the capacity for autonomous choice and critical reflection. For example: individuals may internalise some aspects of their oppression but question others; they may be mistaken about facts, but not values; or they may engage in forced trade-offs amongst well-being in different domains of their lives. Along similar lines Terlazzo emphasises the distinction between global and local autonomy: individuals can have the global capacity for autonomy despite occasional failures to exercise it locally.[45]

Yet despite distinguishing the various ways preferences may be unreliable, these approaches still aim to provide a unifying account, identifying a single central feature that typifies all APs.[46] They make a binary judgement between preferences that are adaptive and unreliable, and those that are not, and this, I argue, obscures the different ways in which preferences may be adaptive.

---

[40] Terlazzo 2015: 10. This substantive procedural approach is similar to one I outline in Begon 2015: 248-249. Specifically, I argue that the valuable options individuals should reflect in light of are central capabilities.
[41] Terlazzo 2015: 11
[42] Terlazzo 2014: 195
[43] Khader 2013: 317-320
[44] Khader 2013: 311
[45] Terlazzo 2015: 12-15. Also on this distinction, Colburn 2010: 4, 21.
[46] This is also true of Barnes's (2016; 2009) account, mentioned above.

2.2 Preferences in a Theory of Justice

On at least some welfarist accounts of distributive justice, no tenable distinction can be drawn between WBAPs and JAPs: if individuals' interests directly determine their entitlements, then if they are a reliable guide to their best interests, they are also a reliable guide to their distributive entitlements (and *vice versa*). This distinction – and, indeed, any discussion of APs at all – will also have no role in, or relevance for, objectivist theories that understand individuals' entitlements (and possibly also their well-being) to be determined entirely independently of their preferences. However, many theories of distributive justice do not fall into either of these camps: whilst individuals' entitlements are not reducible to what would make them happy or satisfied, entitlements are not determined wholly independently of the views and values of those to whom they are provided. In this group we might include Dworkin's equality of resources, Nussbaum's capability approach, Arneson's luck egalitarianism, Cohen's equal access to advantage, and Rawls's theory of justice.[47] (This is not the place to defend a substantive theory of justice. For the sake of clarity I refer to Dworkin's approach, though the general claims apply to any of them.)

Preferences broadly play two roles in such theories of distributive justice. First, they play some role in determining the content of our distributive entitlements. For Dworkin, we are entitled to state assistance for those disadvantages we would consider significant enough to insure ourselves against. Thus, taxation rates are justified by reference to a hypothetical insurance market, the outcome of which depends on our particular tastes and ambitions.[48] Further, state assistance should only be provided if individuals consider themselves to be disadvantaged: if they consider some circumstance a matter of good fortune, we should not compensate them for it.[49] APs can limit the contents of our distributive entitlements, and the realm of state assistance: if we no longer saw mental illness, for example, as something worth insuring against, we would not tax and redistribute to individuals who suffer from it. Further, if an individual does not consider themselves disadvantaged by their condition or impairment, then they would not be entitled to pay-outs (though others with the same condition might be).[50]

[47] E.g. Dworkin 2000; Nussbaum 2000; Arneson 2000; Cohen 1989; Rawls 1999.
[48] Dworkin 2000: 90-109
[49] Dworkin 2000: 294. This is 'the continuity test' (Williams 2002).
[50] In practice, Dworkin may not endorse such an individuated approach, though he acknowledges it is a consequence of this approach (Dworkin 2002; Begon unpub. ms.).

Second, our preferences in a particular case determine whether or not we make use of the opportunities or resources to which we are entitled. Liberal approaches are unlikely to compel individuals to function, both because the intervention required would be illegitimately paternalist, and because an individual's autonomous decision not to exercise an opportunity legitimates its absence. This is a central insight of the capability approach: that a starving, but not a fasting, individual is necessarily a concern of justice. APs here may lead us to wrongly consider self-harm or self-sacrifice to be voluntary, and hence not unjust. I will now consider the various ways in which preferences may be unreliable in either of these roles.

2.3 Justice Adaptive Preferences

WBAPs are procedurally non-autonomous and irrational, since the causes of these preferences are necessarily opaque to their possessors. Some APs do not meet this criterion, yet, nonetheless, should not be used to inform a theory of justice: either to inform our general entitlements, or to justify self-sacrifice as voluntary. These are rational JAPs, and can be divided into three further categories: non-autonomous character planning (where individuals consciously downgrade opportunities they have been wrongly denied), justice adaptive choices (where individuals' counterfactual preferences are autonomous and reliable, but their choices are rendered non-autonomous by the limitations of their circumstances), and the coherent internalisation of oppressive norms (where individuals internalise a self-conception according to which they are not entitled to certain opportunities).

First, non-autonomous character planning. When discussing traditional accounts of APs, I noted that they are usually defined in direct contrast to conscious, rational, and autonomous character planning. Whilst the fox who could not accept that the reason he believes grapes are sour is their inaccessibility is not to be trusted, this is not true of the fox who consciously cultivates a preference for sweeter, lower-hanging fruit. However, this example is misleading. Consciously revising our preferences in light of diminished options *can* be autonomous, but it need not be, depending on the degree to which our options are constrained. For example, someone who has adapted to ableist biases in hiring practices, and concluded that they do not really *need* opportunities for meaningful work, may prioritise personal relationships and hobbies instead. Further, they may be

aware of the reasons for this change of priorities, or would not repudiate them if they were. Such preferences are rational, and do not demonstrate a failure in the individual's autonomous capacities. Indeed, these preferences may function as a reliable guide to their interests (in these unjust circumstances). Nonetheless, we surely should not use them to conclude *either* that the opportunity for meaningful employment is, in general, not a concern of justice, *or* that this individual has voluntarily chosen to forgo this opportunity.

This preference is not autonomous, then, not because the individual is a defective agent, who lacks, or has misused, the capacity for autonomous choice. Rather, their circumstances prevent them from being autonomous: from deciding for themselves what is valuable, and being able to pursue it. Specifically, it seems these circumstances undercut individuals' independence: they are subject to factors that undermine their ability to 'decide for themselves'.[51] Yet in this case, their independence is not undermined by *covert* influences (as §1.2 considered) but by a lack of acceptable alternatives, which the agent may be well aware of.

The second set of rational JAPs can be called 'justice adaptive choices'. When individuals engage in character planning, their underlying preferences change. However, as §1.2 discussed, there are also cases in which we can distinguish what people choose (given limited options) and what they counterfactually prefer. For example, in the earlier case, an individual chose not to partake in higher education whilst preferring that it were available in a more accessible form. Again, individuals' choices in these cases are rational, and reflect a capacity for autonomous choice. Yet, again, their circumstances prevent them from being autonomous.

It may seem that, in this case, individuals do not really have adaptive *preferences* at all, since their counterfactual preferences are autonomous, and reliable from the point of view of justice. This is certainly true, and it is important to acknowledge that when options are constrained, an individuals' choice may only reflect their preferences in a very local sense. Further, that their counterfactual preferences *can* inform a theory of justice: we have no reason to doubt their testimony in determining an account of our general entitlements. Nonetheless, sometimes individuals' choices are the only guidance we have regarding their

[51] For a discussion of independence and its importance for autonomy, see Colburn (2011: 26-31).

preferences, and when they are rendered non-autonomous in this way, they should be deemed justice adaptive: they should not be used to suggest that what they choose to forgo is generally unimportant, or that their lack of it was a voluntary sacrifice. For example, if an individual with a mobility impairment is forced to choose between modifying their home so they can move around it more easily, and pursuing leisure activities, the fact that they choose the former is not evidence that they dispute the value of leisure time in human life, nor that they have voluntarily chosen not to pursue it.

Third, and finally, are cases involving the coherent internalisation of oppressive norms. These are preferences borne out of mistreatment and manipulation so comprehensive that individuals would not repudiate the cause of their preference if they were aware of it, and it does not conflict with their other beliefs and preferences. Consider, for example, if the above sexist science teacher was sufficiently successful that, on being made aware of the reasons we gave up science, we would say the teacher did the right thing, since science, after all, is 'not for women'. As §1.2 considered, these preferences might be a guide to our best interests in light of our preference set. Nonetheless, though this influence on our preferences is not covert (we are not necessarily unconscious of it) it surely violates any plausible interpretation of the independence condition, and so cannot be considered autonomous. As such, these preferences should not determine individuals' distributive entitlements at either the general or individual level.

JAPs need not also be WBAPs, then, but of course they can be. The inaccessibility of grapes may not be a concern of justice, but many preferences that arise as the result of covert influences *are* relevant to justice. WBAPs are also JAPs when the diminished options to which individuals irrationally adapt occur in domains of their life over which they *ought*, as a matter of justice, to be able to exercise autonomy and pursue their own conception of the good. For example, a deaf individuals' preferences not to listen to music and not to be able to engage politically might both be WBAPs, but only the latter is also a JAP. Thus, it should again be emphasised, that whilst we are all likely to have WBAPs, for those in unjust circumstances these are more likely to be JAPs too.

2.4 Responding to JAPs

Different forms of intervention are called for depending on the sense in which individuals' preferences are adaptive. First, when only individuals' *choices* are adaptive (and their counterfactual preferences are reliable), our focus should be providing the acceptable alternatives they lack. Moreover, we should engage in a process of deliberation with the affected individuals to determine what the content of these alternatives should be, since we have no reason to discount their counterfactual preferences. Thus, the state should not intervene in individuals' adaptive choices (on the grounds that they are adaptive choices), but should provide the conditions so that they can choose in a way that reflects their underlying preferences.

However, in at least some cases of non-autonomous character planning, the coherent internalisation of oppressive norms, and irrational JAPs, it is likely that the provision of further options will be insufficient. In these cases, individuals' adaptation to diminished options may have lead to some degree of value distortion: those who have undergone non-autonomous character planning have consciously devalued the opportunities they have been deprived off; those who have undergone the coherent internalisation of oppressive norms have internalised a conception of themselves as not entitled to various opportunities; and those who have irrational JAPs reject certain opportunities for reasons they could not, themselves, accept. I noted two ways in which preferences might play a role in a theory of justice, and value distortions can occur that undermine both these roles: they may affect the content of our general entitlements, and they may cast doubt on whether an instance of self-deprivation is really voluntary.

First, then, an individual may repudiate the value of an opportunity or resource entirely. Such a preference would take the form: 'I live well without x, so x cannot be essential to a decent life'. For example, autistic individuals may question whether certain forms of social interaction are as essential as neurotypical individuals assume, deaf individuals may question whether hearing is essential, and paraplegic individuals may doubt the importance of being able to walk. In determining the content of our distributive entitlements, we may not want to allow individuals who have not experienced some opportunity to veto its inclusion. For example, disabled individuals who reject the value of work, education, or relationships simply because they have been deprived of them as the result of ableist biases should not lead us to conclude that these

opportunities are not a concern of justice. However, it is worth noting that if our just entitlements are conceptualised at a relatively high level of abstraction (as I contend they should be[52]), then individuals who repudiate, for example, the importance of being able to walk or hear do not threaten agreement on their content. The relevant opportunities here are for mobility and aesthetic experience, and these individuals have not been deprived of these capabilities, nor do they repudiate their value.

Moreover, individuals with APs do not tend to reject the importance of some opportunity in general, they merely repudiate its value in their life. Such preferences might take the form: 'x might be good for others, but it's not for people like me – and I'm happy without it'. For example, an individual with Down's syndrome might take the view that education is valuable, but conclude that 'it's not for them' if their educative environment is not an inclusive one. These individuals' preferences can be used to determine the opportunities that are relevant to justice, but should not be used as evidence that an individual has what they are entitled to: the individual with Down's does not dispute the value of education, but their 'choice' not to engage in it should not lead us to conclude that justice has been done. A voluntary choice to reject an option requires not just that acceptable alternatives are available, but that these choices are considered live options. Individuals need, in Nussbaum's terminology, the 'internal capability' to exercise an opportunity,[53] as well as the mere freedom to do so. Exactly how these meaningful opportunities can be provided is a complex question, and will likely require some form of deliberation with affected individuals, which involves consciousness-raising (such that they understand the significance and potential value of certain options) as well as the provision of opportunities.

Justice adaptive choices, then, can be avoided by providing individuals with options that reflect their preferences, where they are entitled to these as a matter of justice. Other JAPs arise from value distortions, and avoiding these requires both that individuals have acceptable options open to them, so their values are not distorted by a lack of exposure to reasonable alternatives, *and* being taught to see why these options might be valuable and to see themselves as the kind of people who can exercise them.

[52] Begon 2017
[53] E.g. Nussbaum 2011: 20-23

Finally, it should be emphasised that, though APs are non-autonomous, diagnosing someone as having one need not be insulting.[54] First, having rational JAPs does not indicate any lack of the internal capacities for autonomy: the problem lies in circumstances. Second, having a value distortion on some particular issue does not imply an individual is wholly unreliable, or lacks the capacity to make autonomous choices or form reliable preferences in any area of their life. Finally, though irrational JAPs do indicate flawed reasoning, it does not follow that their possessors are *unusually* irrational or incapable of formulating consistent preferences. We all adapt to restrictions in our options, and many of us do so in non-autonomous ways (we have WBAPs). However, when these restrictions stem from oppression and mistreatment, the resulting adaptations are more likely to be a concern of justice (JAPs). The fault, though, lies in circumstances rather than individuals. The reason individuals in unjust circumstances are more prone to JAPs is not that they are more defective agents.

3. Conclusion

I began this paper with a dilemma that discussions of AP raise: on the one hand, we do not want to disregard the preferences of oppressed groups, and ignore their members as untrustworthy agents, yet on the other, we do not want to allow that cases in which individuals have adapted to deprivation no longer constitute an injustice. I have argued that it is only JAPs that should be considered unreliable in the context of social justice, and that JAPs need not be irrational. When disabled individuals' preferences are JAPs, this is likely to be the result of their circumstances, which, insofar as this is possible, should be changed. We have no reason to think disabled individuals are unusually deficient in autonomy, or more prone than others to adapting to their circumstances, and we should not dismiss the preferences of any individual or group out of hand. However, we are in danger of perpetuating injustice if we fail to acknowledge the ways in which restrictions of individuals' options can undermine their ability to form autonomous preferences in areas relevant to justice, and prevent them from effectively pursuing those options they prefer, and should be entitled to. Thus, when preferences are justice adaptive, they should not determine the contents of our distributive entitlements, nor determine whether a sacrifice is voluntary.

[54] I do not mean to downplay individuals' experiences, and do not deny that (especially given currently unjust social circumstances) this may be experienced as insulting.

References

Albrecht, G.L. and Devlieger, P.G. (1999) 'The Disability Paradox: High Quality of Life Against All the Odds', *Social Science and Medicine* 48, 8: 977-988.

Arneson, R.J. (2000) 'Luck Egalitarianism and Prioritarianism', *Ethics* 110, 2: 339-349

Baber, H.E. (2007) 'Adaptive Preferences', *Social Theory and Practice* 33, 1: 105-126

Barnes, E. (2016) *The Minority Body: A Theory of Disability* (Oxford: Oxford University Press)

Barnes, E. (2009) 'Disability and Adaptive Preference', *Philosophical Perspectives* 23: 1-22

Begon, J. (2017) 'Capabilities for All? From Capabilities to Function, to Capabilities to Control', *Social Theory and Practice* 43, 1: 154-179

Begon, J. (2015) 'What are Adaptive Preferences? Exclusion and Disability in the Capability Approach', *Journal of Applied Philosophy* 32, 3: 241-257

Begon, J. (unpub. ms.) 'Disadvantage, Disagreement, and Disability: Re-evaluating the Continuity Test'

Bovens, L. (2002) 'Sour Grapes and Character Planning', *The Journal of Philosophy* 86, 2: 57-78

Cohen, G.A. (1989) 'On the Currency of Egalitarian Justice', *Ethics* 99: 906-944

Colburn, B. (2011) 'Autonomy and Adaptive Preferences', *Utilitas* 23, 1: 52-71

Colburn, B. (2010) *Autonomy and Liberalism* (New York: Routledge)

Dworkin, R. (2002) '*Sovereign Virtue* Revisited', *Ethics* 113, 1: 106-143.

Dworkin, R. (2000) *Sovereign Virtue* (Cambridge, MA: Harvard University Press)

Elster, J. (1987) *Sour Grapes: Studies in the Subversion of Rationality*, (Cambridge: Cambridge University Press)

Jaggar, A. (2005) 'Saving Amina: Global Justice for Women and Intercultural Dialogue', *Ethics and International Affairs* 19: 55–75

Khader, S. (2013) 'Identifying Adaptive Preferences in Practice: Lessons from Postcolonial Feminisms', *Journal of Global Ethics* 9, 3: 311-327

Khader, S. (2012) 'Must Theorising about Adaptive Preferences Deny Women's Agency?', *Journal of Applied Philosophy* 29, 4: 302-317

Khader, S. (2011) *Adaptive Preferences and Women's Empowerment* (Oxford: Oxford University Press)

Khader, S. (2009) 'Adaptive Preferences and Procedural Autonomy', *Journal of Human Development and Capabilities* 10, 2: 169-187

Narayan, U. (2002) 'Minds of Their Own: Choices, Autonomy, Cultural Practices, and Other Women', in L. Antony & C. Witt (eds) *A Mind of One's Own* (Boulder, CO: Westview Press), pp. 418–433

Nussbaum, M.C. (2011) *Creating Capabilities*, (Cambridge, MA: Harvard University Press)

Nussbaum, M.C. (2006) *Frontiers of Justice*, (Cambridge, MA: Harvard University Press)

Nussbaum, M.C. (2004) 'On Hearing Women's Voices: A Reply to Susan Okin', *Philosophy & Public Affairs* 32, 2: 193-205

Nussbaum, M.C. (2001) 'Symposium on Amartya Sen's philosophy: 5 Adaptive preferences and women's options', *Economics and Philosophy* 17, 1: 67-88

Nussbaum, M.C. (2000) *Women and Human Development: The Capabilities Approach* (Cambridge: Cambridge University Press)

Qizilbash, M. (2007) 'Well-Being, Adaptation and Human Limitations', in S. Olsaretti (ed.) *Preferences and Well-Being* (Cambridge: Cambridge University Press) pp.83-109

Rawls, J. (1999) *A Theory of Justice* (Cambridge, MA: Harvard University Press)

Sen, A. (1999) *Commodities and Capabilities* (Oxford: Oxford University Press)

Sen, A. (1988) *On Ethics and Economics* (Malden, MA: Blackwell)

Terlazzo, R. (2015) 'Conceptualizing Adaptive Preferences Respectfully: An Indirectly Substantive Account', *Journal of Political Philosophy* 24, 2: 1-22

Terlazzo, R. (2014) 'The Perfectionism of Nussbaum's Adaptive Preferences', *Journal of Global Ethics*, 10, 2: 183-198

Williams, A. (2002) 'Dworkin on Capability', *Ethics* 113, 1: 23-39

Zimmerman, D. (2003) 'Sour Grapes, Self-abnegation and Character Building', *The Monist* 86: 220-241